**Claims**

1.    A method of processing free-format data stored in a computing system, comprising the steps of examining elements of the data to determine attributes of the data,

5    by examining the content of the elements and the contextual relationships of elements to each other, to determine semantic and syntactic information (attributes) about the data, producing additional data relating to this information, in the form of a text object which includes

10    pointer means enabling access to the elements of the free-format data, and the additional data being accessible by a query processing means to provide answers to queries relating to the semantic and syntactic information about the data and/or to access the data to manipulate the data.

15    2.    A method in accordance with claim 1, wherein the free-format data is stored as a record in a free-format field of a database.

3.    A method in accordance with claim 1 or claim 2, wherein the data remains stored in the computing system as

20    it was originally stored, whereby it may be accessed by other applications.

4.    A method in accordance with any preceding claim, wherein the text object includes an attribute - type identifier which identifies an attribute type of an element

25    of the data.

5.    A method in accordance with any preceding claim, wherein the text object includes a value indicating the character length of an element of the data.

6.    A method in accordance with claim 4 or claim 5,

30    wherein the text object includes a value indicating whether an element is low level in a syntactic hierarchy or higher level whereby the value may be used for matching purposes when matching data with other data processed in accordance with the method.

35    7.    A method in accordance with any preceding claim,

the text object including a match weighting value for an element of the data, which can be used to determine the significance of the element when matching with other free format data.

5      8.      A method in accordance with any preceding claim, wherein the text object comprises a plurality of component nodes arranged according to the semantic structure of the free-format data, the component nodes being arranged in a hierarchy corresponding to the semantic structure of the

10     free-format data and each component node including additional data relating to the corresponding element of the free-format data.

9.      A method in accordance with any preceding claim, comprising the further step of generating matching values

15     for comparing an element of the free-format data with an element of other free-format data processed in accordance with the present method.

10.     A method in accordance with claim 9 where the matching value is a phonetic value for phonetically

20     comparing elements of free-format data.

11.     A method in accordance with any preceding claim, wherein the text object includes implied data relating to information implied from the free-format data.

12.     A method in accordance with any preceding claim,

25     wherein a plurality of free-format data records are processed and a text object associated with each free-format data record is produced.

13.     A method in accordance with claim 12, wherein the text object is stored in the computer system whereby it is

30     available for queries on the associated free-format data record via the query processing means.

14.     A method in accordance with claim 12 comprising the further step of producing a text object index including attribute type identifiers for elements of each data record

35     and pointers to each data record, whereby the index may be

queried by queries relating to semantic and syntactic information about the data and the data may be accessed via the index.

15. A method in accordance with claim 14 wherein each entry in the text object index includes a representative value key, which gives a value representative of a feature of the element associated with the attribute - type identifier.

16. A method in accordance with any preceding claim, comprising the further step of carrying out a domain construction process to construct a domain object from domain definition data files, the domain object being arranged to carry out the examination process by parsing the free-format data in accordance with grammar rules.

17. A method in accordance with claim 16, wherein the domain definition data files include character definition data, regular expression definition data and grammar data.

18. A method in accordance with any preceding claim, wherein the free-format data is postal address data.

19. A method in accordance with any preceding claim wherein the query processing means can carry out normal database operations on the data via the additional data.

20. A processing system for processing free-format data stored in a computing system, the apparatus including means for examining elements of the data to determine attributes of the data, by examining the content of the elements and the contextual relationships of elements to each other, to determine semantic and syntactic information (attributes) about the data, means for producing additional data relating to this information, in the form of a text object which includes pointer means enabling access to the elements of the free-format data, and a query processing means which is arranged to access the additional data to provide answers to queries relating to the semantic and syntactic information about the data and/or to access the

data to manipulate the data.

21. A processing system in accordance with claim 20, wherein the free-format data is stored as a record in a free-format field of a database.

5 22. A processing system in accordance with claim 20 or claim 21, wherein the examining means does not affect the storage of the data.

23. A processing system in accordance with any one of claims 20 to 22, wherein the text object includes an

10 attribute - type identifier which identifies an attribute type of an element of the data.

24. A processing system in accordance with any one of claims 20 to 23, wherein the text object includes a value indicating the character length of an element of the data.

15 25. A processing system in accordance with claim 23 or claim 24, wherein the text object includes a value, indicating whether an attribute - type of an element is low level in a syntactic hierarchy or high level whereby the value may be used for matching purposes when matching with

20 other free-format data processed in accordance with this system.

26. A processing system in accordance with any one of claims 20 to 25, wherein the text object includes a match weighting value for an element of the data, which can be

25 used to determine the significance of the element when matching with other free-format data.

27. A processing system in accordance with any one of claims 20 to 26, wherein the text object comprises a plurality of component nodes arranged according to the

30 semantic structure of the free-format data, the component nodes being arranged in a hierarchy corresponding to the semantic structure of the free-format data, and each component node including additional data relating to the corresponding element of free-format data.

35 28. A processing system in accordance with any one of

claims 20 to 27, the text object means for generating matching values for comparing an element of the free-format data with an element of other free-format data processed by the processing system.

5      29.   A processing system in accordance with claim 28, wherein the matching value is a phonetic value for phonetically comparing elements of free-format data.

       30.   A processing system in accordance with any one of claims 20 to 29, wherein the text object includes implied

10     data relating to information implied from the free-format data.

       31.   A processing system in accordance with any one of claims 20 to 30, wherein the system is arranged to process a plurality of free-format data records and produce a text

15     object associated with each free-format data record.

       32.   A processing system in accordance with claim 31, wherein the means for producing additional data is arranged to produce a text object index including attribute - type identifiers for elements of each data record and pointers

20     to each data record and wherein the query processing means is arranged to access the text object index to provide answers to queries relating to the semantic and syntactic information about the data and/or to access the data to manipulate the data.

25     33.   A processing system in accordance with claim 32, wherein the text object index includes representative value keys for entries, which give a value representative of a feature of the element associated with the attribute - type identifier for the entry for facilitating matching with

30     other free-format data processed in accordance with this system.

       34.   A processing system in accordance with any one of claims 20 to 33, further comprising a domain object, the domain object being arranged to carry out the examination

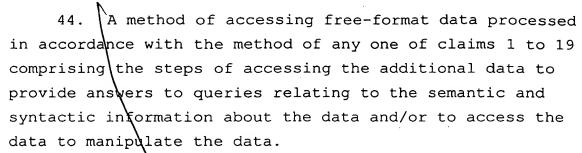35     process by parsing the free-format data in accordance with

grammar rules.

35. A processing system in accordance with claim 34, wherein the domain object is produced by a domain construction process from domain definition data files.

36. A processing system in accordance with claim 35, further comprising a domain constructor for carrying out the domain construction process.

37. A processing system in accordance with claim 35 or claim 36, wherein the domain definition data files include character definition data, regular expression definition data and grammar data.

38. A processing system in accordance with any one of claims 20 to 37, wherein the free-format data is postal address data.

39. A processing system in accordance with any one of claims 20 to 38, wherein the query processing means is arranged to carry out normal database operations on the data via the additional data.

40. A method of enabling access to free-format data stored in a computing system, including a plurality of free-format data records, comprising the steps of storing additional data relating to semantic and syntactic information (attributes) about the data for each data record, the additional data being in the form of a text object associated with each data record, the text object including pointer means enabling access to elements of each free-format data record, the additional data being accessible by a query processing means to provide answers to queries relating to the semantic and syntactic information about the data and/or to access the data to manipulate the data.

41. A processing system for enabling access to free-format data stored in a computing system, including a plurality of free-format data records, the processing system comprising additional data relating to semantic and

syntactic information (attributes) about the data for each
data record, stored and accessible by the processing
system, the additional data being in the form of a text
object associated with each data record, the text object
5 including pointer means enabling access to elements of each
free-format data record, and a query processing means
arranged to access the additional data to provide answers
to queries relating to the semantic and syntactic
information about the data and/or to access the data to
10 manipulate the data.

42. A method of enabling access to free-format data
stored in a computing system, including a plurality of
free-format data records, comprising the steps of storing
additional data relating to semantic and syntactic
15 information (attributes) about the data of each data
record, the additional data being in the form of a text
object index which includes attribute - type identifiers
for elements of each data record and pointers to each data
record, the text object index being accessible by a query
20 processing means to provide answers to queries relating to
the semantic and syntactic information about the data
and/or to access the data to manipulate the data.

43. A processing system for enabling access to
free-format data stored in a computing system, including a
25 plurality of free-format data records, the processing
system comprising the additional data relating to semantic
and syntactic information (attributes) about the
free-format data for each data record, the additional data
being in the form of a text object index which includes
30 attribute type identifiers for elements of each data record
and pointers to each data record, and a query processing
means arranged to access the additional data to provide
answers to queries relating to the semantic and syntactic
information about the data and/or to access the data to
35 manipulate the data.

44. A method of accessing free-format data processed in accordance with the method of any one of claims 1 to 19 comprising the steps of accessing the additional data to provide answers to queries relating to the semantic and

5    syntactic information about the data and/or to access the data to manipulate the data.

45. A processing system for enabling access to free-format data processed in accordance with the method of any one of claims 1 to 19, the processing system including

10    a query processing means arranged to access the additional data and provide answers to queries relating to the semantic and syntactic information about the data and/or to access the data to manipulate the data.

46. A processing system for processing free-format

15    data stored in a computing system, comprising means for examining elements of the data to determine attributes of the data, by examining the content of the elements and the contextual relationship of elements to each other, to determine semantic and syntactic information (attributes)

20    about the data, and a query processing means for utilising this information to provide answers to queries relating to the semantic and syntactic information about the data and/or to access the data.

47. A processing system in accordance with claim 46,

25    wherein the examining means retains the free-format data as stored in the computer system, without affecting it.

48. A method of processing free-format data stored in a computing system, comprising the steps of examining elements of the data to determine attributes of the data,

30    by examining the content of the elements and the contextual relationships of elements to each other, to determine semantic and syntactic information (attributes) about the data, and querying the data using this information to provide answers to queries relating to the semantic and

35    syntactic information about the data and/or to access the

data.

49. A method of processing free-format data in accordance with claim 48, wherein the free-format data is unaffected by the examining process and remains stored in
5      the computing system as it was originally stored.

50. A computer readable memory storing instructions for controlling a computer to process free-format data stored in a computing system, in accordance with the method of any one of claims 1 to 19.

10      51. A computer readable memory storing instructions for controlling a computer to process free-format data stored in a computing system, in accordance with the method of claim 48.

52. A method of processing a plurality of records of
15      free-format data stored in a computing system, comprising the steps of, for each record, examining elements of the data to determine attributes of the data, by examining the content of the elements and the contextual relationships of elements to each other, to determine semantic and syntactic
20      information (attributes) about each record, and producing virtual data fields associated with each record enabling access to this information and the associated elements, whereby each record is provided with associated virtual data fields enabling access to semantic and syntactic
25      information about that record and also access to the associated elements.

53. A processing system for processing free-format data records stored in a computing system, comprising means for examining elements of the data of each record to
30      determine attributes of the data, by examining the content of the elements and the contextual relationship of elements to each other, to determine semantic and syntactic information (attributes) about the data, and means for producing virtual data fields associated with each record
35      enabling access to this information and the associated

elements, whereby each record is provided with associated
virtual data fields enabling access to semantic and
syntactic information about that record and also access to
the associated elements.

5

add
A₁

add
B₂